UNR Joint Economics Working Paper Series
Working Paper No. 07-014

Intercept and Recall: Examining Avidity Carryover in On-Site Collected Travel Data

Klaus Moeltner and J. Scott Shonkweiler

Department of Resource Economics / MS 204
University of Nevada, Reno
Reno, NV 89557
(775) 784-6701 | Fax (775) 784-1342
email: moeltner@cabnr.unr.edu

December, 2007

Abstract

This study examines the proper estimation of trip demand and economic benefits for visitors to recreation sites when past-season trip information is elicited from travelers intercepted on-site. We show that the proper weighting of past season counts is different from the standard on-site correction appropriate for current-season counts. We find that for our sample of lake visitors relatively stronger preference or "avidity" for the interview site carries over across seasons. We further show that using the correct weighting of past trip counts is critical in deriving meaningful estimates of travel demand and economic benefits.

JEL Classification: C35, C51, Q26

Keywords: On-site Sampling; Recreation Demand Systems; Poisson-Lognormal Distribution; Simulated Maximum Likelihood

Intercept and Recall:
Examining Avidity Carryover in On-Site Collected Travel Data

Klaus Moeltner* J. Scott Shonkwiler

Draft December 7, 2007

*Contact author

Both authors:
Department of Resource Economics
MS 204
University of Nevada, Reno
Reno, Nevada 89557
Phone: (775) 784-4803
Fax: (775) 784-1342
e-mail: moeltner@cabnr.unr.edu

Intercept and Recall:
Examining Avidity Carryover in On-Site Collected Travel Data

Abstract:

This study examines the proper estimation of trip demand and economic benefits for visitors to recreation sites when past-season trip information is elicited from travelers intercepted on-site. We show that the proper weighting of past season counts is different from the standard on-site correction appropriate for current-season counts. We find that for our sample of lake visitors relatively stronger preference or "avidity" for the interview site carries over across seasons. We further show that using the correct weighting of past trip counts is critical in deriving meaningful estimates of travel demand and economic benefits.

JEL Codes: C35, C51, Q26

Key words: On-site Sampling, Recreation Demand Systems, Poisson-Lognormal Distribution, Simulated Maximum Likelihood

Introduction

Two important characteristics of destinations for recreational travel are accessibility and the type and quality of on-site amenities. To assure an adequate provision of these public goods, agencies need to maintain access roads and on-site infrastructure. For sites gaining in popularity, investments that enhance the capacity and quality of roads may be required in order to provide dependable access, to reduce congestion, or to protect environmentally sensitive areas from vehicular encroachment.

Naturally, from a social efficiency perspective these investments ought to be commensurate to the economic gains enjoyed by the users of a given site. These welfare gains are theoretically conceived as monetary equivalents of upward changes in an individual's utility function, and traditionally interpreted as "willingness-to-pay" (WTP) to enjoy a given site visit or a series of visits over an entire season. The knowledge of site-specific welfare effects is also important when considering the optimal length and timing of planned road closures, as they may be required for maintenance operations and capacity expansions. While pure traffic volume associated with a given destination can be measured by automatic devices such as electronic vehicle counters, the assessment of economic benefits requires surveying visitors.

An important input to the estimation of such welfare effects is information on the total number of trips per season made by visitors. General population surveys of recreation activities are difficult to implement, may have poor response rates, and are generally much more costly to administer than surveys conducted at a recreational site. So we consider sampling via direct interaction with visitors intercepted at the interview site and possibly other nearby sites that can be considered close substitutes. This raises the issue of the appropriate timing of such on-site polling. Specifically, any sampling before the end of a visitation season or planning year imposes the risk of forecasting errors on part of the respondent when asked to report the estimated total number of trips for the current season. These errors can be especially large for sites with pronounced quality changes during a given season. For example, as is the case for the application underlying this study, water levels at reservoir destinations may change dramatically even over the span of a few weeks. This may hamper motorized water sports, but possibly enhance beach recreation or fishing access. Another example would be unexpected weather events that may affect site quality or shorten a recreational season, especially at higher elevations.

To avoid forecasting errors the researcher may thus opt for end-of-season sampling. However, this carries the risk of sample selection if end-of-season users are systematically different from other users in the wider population of visitors. Intercepting visitors throughout the season is likely to generate a more representative sample of the underlying population of recreationists. Therefore, if the researcher has strong concerns regarding the accuracy of forward-looking trip reports but at the same time desires to sample visitors throughout the season to avoid selection problems and to assure a reasonable sample size within the limits imposed by constrained survey budgets, an attractive option might be to ask respondents about visits during *past seasons* in addition to or instead of projected visits for the current or future seasons.[1] This information on past visits can then form the basis for the assessment of current and future travel demand and the estimation of economic welfare gains.

This study focuses on the appropriate statistical framework to process such data. Ultimately the researcher desires to derive welfare estimates that apply to the “prototypical” individual in the underlying population of visitors. It has been empirically shown numerous times that on-site collected trip reports for the interview site are generally inflated compared to the number of visits to that site taken by the typical user. This discrepancy arises because intercepted respondents are likely to have a relatively stronger preference for the examined site than the prototypical user in the underlying population. The enhanced avidity of the respondent for the site of interception must be explicitly captured in the modeling framework to avoid biased trip and welfare estimates (e.g. Englin and Shonkwiler, 1995, Moeltner and Shonkwiler, 2005).

It is by now well understood how to implement statistical corrections for on-site trip counts associated with the *current season* (Englin and Shonkwiler, 1995, Moeltner and Shonkwiler, 2005, Egan and Herriges, 2006). The question now arises if this "avidity" correction is also necessary for trip reports associated with past seasons. On the one extreme if recreational travelers randomly re-order their preferences for different destinations at the beginning of each season, the intensity of past season visits for the current interview site will have no systematic bearing for preferences for the same site this season. At the other extreme, if site preferences are perfectly stable over time a full-fledged on-site correction as has traditionally been employed for current season counts may be required. The first scenario of random preference-reshuffling appears somewhat unlikely for the prototypical on-site respondent, especially in absence of pronounced changes in site quality or availability of substitute sites over adjacent seasons. The second scenario of "complete avidity-carryover" is equally questionable as it rules out even subtle changes in user preferences or any kind of variety-seeking behavior (e.g. Moeltner, 2004).

The aim of this study is to derive a flexible statistical estimator for processing on-site collected information on past trips that accommodates both of these extremes, but also allows for the more likely outcome of "attenuated avidity carryover". While our proposed model is applicable to any type of recreation destination or activity we focus in this study on jet skiers intercepted at various High Sierra lakes in the Tahoe Region. For this sample we find that avidity carryover is very pronounced, but far from complete. Using the appropriate econometric model has substantial implications for the estimation of travel demand and economic welfare.

The remainder of this manuscript is structured as follows: The next section develops the theoretical and econometric framework for a multi-site travel demand model based on past season counts. Section III presents the empirical application, including the description of our data, a discussion of estimation results, and the derivation of trip and welfare predictions. Concluding remarks are given in Section IV.

II) Model Formulation

Utility-theoretic Framework

We stipulate that person *i* derives aggregate utility in season *t* from taking trips to the *j=1...J*-site recreation system collected in vector $\mathbf{y_{it}}$ and from consuming a numeraire composite commodity *b*. Specifically,

$$U_{it} = U\left(\mathbf{y_{it}}, \mathbf{q_t}, \mathbf{h_i}, b\right), \tag{1}$$

where $\mathbf{q_t}$ denotes site attributes, and $\mathbf{h_i}$ is a vector of person or household characteristics. Assuming seasonal separability of utility[2], we apply the Incomplete Demand System (IDS) framework described in LaFrance and Hanemann (1989). Utility maximization subject to an (assumed binding) budget constraint yields the Marshallian quasi-demand system

$$\mathbf{y_{it}} = \mathbf{y}\left(\mathbf{p_i}, \mathbf{q_t}, \mathbf{h_i}, m_i\right) \tag{2}$$

where $\mathbf{p_i}$ is a vector of prices associated with the destinations included in the system, and m_i denotes annual income.[3] As shown in LaFrance and Hanemann (1989) these demand equations display, in theory, all desired utility-theoretic properties. LaFrance and Hanemann (1994) illustrate how this framework can be empirically implemented for some common functional forms of demands. We follow Shonkwiler (1999) and apply a Log II demand specification within a count data framework. We initially specify trips to follow a Poisson distribution with expected site-specific demand given by

$$E\left[y_{ijt}\right] = \lambda_{ijt} = \exp\left(\mathbf{a_{ijt}} + \beta_{p,j} p_{ij}\right) \cdot m_i^{\beta_m} \quad , \tag{3}$$

where we have implemented the utility-theoretic IDS restrictions on price coefficients of $\beta_{p,ik} = 0, k \neq j$, and $\beta_{m,i} = \beta_m, \forall i$,[4] and simplifying restriction $\beta_{p,ij} = \beta_{p,j}\ \forall i$. Shifting vector $\mathbf{a_{ijt}}$ comprises all site and respondent characteristics multiplied by their respective coefficients.

Unobserved Heterogeneity

It is likely that trip demand includes respondent-specific components that are unobserved by the researcher. This individual heterogeneity can be conveniently incorporated into a Poisson system by combining expected demand for the Poisson distribution with a multiplicative lognormal error term. Heterogeneity-adjusted conditional expected demand is thus given by

$$E\left[y_{ijt} \mid \varepsilon_{ij}\right] = \lambda_{ijt} \cdot \exp\left(\varepsilon_{ij}\right) = \tilde{\lambda}_{ijt} \tag{4}$$

The choice of a normal distribution for ε_{ij} lead's to Aitchison and Ho's (1989) Poisson-lognormal model. This specification was first implemented in the recreation demand context by Shonkwiler (1995), and recently examined in more detail in the context of on-site sampling by Egan and Herriges (2006). As described in that study the Poisson-lognormal distribution has several desirable properties for the estimation of recreation demand to a system of sites. We will thus adopt this density for our application.

To keep our model tractable we assume that error terms are individual and site-specific, but invariant over time. An intuitive interpretation of this error specification would be the presence of unobserved individual preferences for an equally unobserved (or costly-to-measure) site-specific quality (e.g. choice and difficulty of nearby hiking trails, water depth at nearby beaches etc.) and an abundance or lack of such quality at site *j*.

To link trip reports associated with a given respondent across sites we specify that the heterogeneity terms, $\varepsilon_{ij},\ j = 1 \cdots J$, follow a multivariate normal distribution with full variance-covariance matrix, i.e.

$$\boldsymbol{\varepsilon}_{\mathbf{i}} = \begin{bmatrix} \varepsilon_{i1} & \varepsilon_{i2} & \cdots & \varepsilon_{iJ} \end{bmatrix}' \sim mvn(\mathbf{0}, \boldsymbol{\Sigma}) \quad \text{with}$$
$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1J} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{J1} & \sigma_{J2} & \cdots & \sigma_{JJ} \end{bmatrix} \quad \text{and} \quad \mathrm{E}\left[\boldsymbol{\varepsilon}_{\mathbf{i}} \boldsymbol{\varepsilon}_{\mathbf{g}}'\right] = 0\ g \neq i. \tag{5}$$

As indicated in (6), all individuals share a common variance-covariance matrix for the heterogeneity error vector.[5]

The unconditional density of $\mathbf{y}_{\mathbf{it}}$ is thus Poisson-lognormal (PLN, (Aitchison and Ho, 1989)) with

$$f(\mathbf{y}_{\mathbf{it}}) = \int_{\boldsymbol{\varepsilon}_{\mathbf{i}}} \prod_{j=1}^{J} \frac{\exp\left(-\tilde{\lambda}_{ijt}\right) \cdot \tilde{\lambda}_{ijt}^{\ y_{ijt}}}{y_{ijt}!} f(\boldsymbol{\varepsilon}_{\mathbf{i}})\, d\,\boldsymbol{\varepsilon}_{\mathbf{i}}, \tag{6}$$

where $f(\boldsymbol{\varepsilon}_{\mathbf{i}})$ denotes the multivariate normal density, and the dimensionality of the integral is commensurate to the number of elements in $\boldsymbol{\varepsilon}_{\mathbf{i}}$. The desirable properties of this mixture distribution in the context of recreation demand systems are discussed in detail in Egan and Herriges (2006). Borrowing from their notation, the first two moments of the unconditional marginal distribution can be derived as

$$E\left[y_{ijt}\right] = \lambda_{ijt} \cdot \exp\left(\frac{\sigma_{jj}}{2}\right) = \delta_{ijt} \quad \text{and}$$
$$\mathrm{V}\left[y_{ijt}\right] = \delta_{ijt} + \delta_{ijt}^{2} \cdot \left(\exp\left(0.5 \cdot \sigma_{jj}\right) - 1\right). \tag{7}$$

Correcting for on-site sampling

As originally discussed in Patil and Rao, (1978), if the population density of a random variable *x* is given by $f(x)$, the weighted or "size-biased" density for the same variable measured on-site takes the form of $f^{s}(x) = (x/E[x]) \cdot f(x)$. The proper statistical approach for addressing on-site sampling for the univariate Poisson distribution is shown in Shaw (1988). Egan and Herriges (2006) extend this framework to the Poisson-lognormal distribution and show how the density in (6) can be corrected for *current* season counts that are collected on-site. Specifically, if the interview is conducted at site *k*, the joint trip density for the current season to the system of sites takes the form

$$f^{k}\left(\mathbf{y_{it}}\right) = \frac{y_{ikt}}{E\left[y_{ikt}\right]} \int_{\boldsymbol{\varepsilon}_{\mathbf{i}}} \prod_{j=1}^{J} \frac{\exp\left(-\tilde{\lambda}_{ijt}\right) \cdot \tilde{\lambda}_{ijt}^{\,y_{ijt}}}{y_{ijt}!} f\left(\boldsymbol{\varepsilon}_{\mathbf{i}}\right) d\,\boldsymbol{\varepsilon}_{\mathbf{i}} \tag{8}$$

where the superscript to *f(.)* indexes the intercept site. The term outside the integral is the multiplicative weight assigned to the density of y_{ikt}.

This raises the question if past season counts for the current interview site also need to be weighted to allow for pronounced avidity for the intercepted respondent with respect to the interview destination. We approach this as an empirical issue and propose what Patil and Rao (1978) deem a general weighting function that, in our context, accommodates the extreme cases of "zero avidity carryover" and "complete avidity carryover". Specifically, using subscript *t-1* to denote the season preceding the sampling season, we specify the size-corrected joint Poisson-lognormal density as

$$f^{k}\left(\mathbf{y_{it-1}}\right)=\frac{w\left(y_{ikt-1}\right)}{E\left[w\left(y_{ikt-1}\right)\right]}\int_{\boldsymbol{\varepsilon}_\mathbf{i}}\prod_{j=1}^{J}\frac{\exp\left(-\tilde{\lambda}_{ijt}\right)\cdot\tilde{\lambda}_{ijt}^{\;y_{ijt}}}{y_{ijt}!}f\left(\boldsymbol{\varepsilon}_\mathbf{i}\right)d\,\boldsymbol{\varepsilon}_\mathbf{i} \qquad \text{where}$$
$$w\left(y_{ikt-1}\right)=\left(y_{ikt-1}+1\right)^{\alpha}-\alpha, \qquad 0\le\alpha\le1. \tag{9}$$

The weighting function *w(.)* has several desirable properties. First, it accommodates counts of zero and thus a scenario where a respondent intercepted at site *k* in the current season did not visit that site in the preceding year. Second, as α approaches zero the weight term in (9) approaches one. This would imply that no on-site correction is needed for past counts. We label this outcome as "zero carryover". Third, as α approaches one the weight term takes the form of the weight in (8), indicating that a full on-site correction is needed for past season visits to the interview site, i.e. that there is "complete carryover" of avidity for the intercept site. A value of α between zero and one would imply that past season trips to the interview site require a different on-site correction than current season counts. Naturally, a shortcoming of this formulation is that *w(.)* equals zero for the joint outcome of $\alpha=1$ and $y_{ikt-1}=0$. In that case, the weighting term in (9) is no longer well defined as its denominator goes to zero. Since there are a few individuals in our data that reported zero trips to the interview site for the preceding season, we cannot directly implement a restricted version of the specification in (9) with an imposed constraint of $\alpha=1$ to formally test for "full carryover". Instead, as shown in the estimation section, we use the empirical confidence interval for α to examine this hypothesis.

For this study, we estimate both unweighted model (6) and weighted model (9) using counts for the season preceding the sampling year. The integrals in (6) and (9) are simulated using 1000 draws of Halton vectors (e.g. Train, 1999). The computation of the sampling weight in (9) poses somewhat of a challenge as its denominator depends on $\boldsymbol{\varepsilon}_\mathbf{i}$ and does not have closed form. Specifically,

$$E\left[w\left(y_{ikt-1}\right)\right]=\sum_{y_{ikt-1}=0}^{\infty}\left(w\left(y_{ikt-1}\right)\cdot f\left(y_{ikt-1}\right)\right)=\sum_{y_{ikt-1}=0}^{\infty}\left(w\left(y_{ikt-1}\right)\cdot\int_{\varepsilon_{ik}} f\left(y_{ikt-1}\mid\varepsilon_{ik}\right) f\left(\varepsilon_{ik}\right) d\varepsilon_{ik}\right) \qquad (10)$$

where $f\left(y_{ikt-1}\right)$ is the marginal Poisson-lognormal density of y_{ikt-1}. The integral in the last term in (10) must be simulated apart from the simulation routine used to evaluate the integral in (9). In addition, the summation over $y_{ik,t-1}$ in (10) needs to be numerically approximated as well.[6] Our general estimation framework is simulated maximum likelihood (e.g. Train, 2003). The algorithm produces estimates of the slope coefficients in (3), the elements of variance-covariance matrix $\mathbf{\Sigma}$, and, for the past season model, the avidity parameter α.[7]

III) Empirical Application

Data

The data for this analysis stem from an on-site survey of jet skiers implemented during the summer seasons of 2001 and 2002 at six lakes and reservoirs in the Tahoe region of the central Sierra Nevada. A detailed description of the survey procedures is provided in Moeltner and Shonkwiler (2005). For this study we use information on past season trips to five of the six lakes[8]. To be specific, visitors interviewed in 2001 provided trip information for the years 2000 and 2001. A different set of respondents, captured in the 2002 round of the survey, reported trips for 2001 and 2002. After eliminating individuals who took more than 40 trips to the system of sites and / or spent more than one day at the interview location[9], we retain 159 completed questionnaires yielding a panel of 159x5 = 795 observations for "past year trips" (= trips in 2000 for the 2001 sample, and trips in 2001 for the 2002 sample).[10]

Table 1 summarizes some basic lake and trip characteristics for this sample. As can be seen from the table visitors reported a total of 1029 trips to the recreation system during the season preceding the interview year. The largest numbers of seasonal trips are observed for Lahontan and Boca reservoirs. Both destinations offer numerous easy access and launching points, generally free of charge. Distances from visitor origin to destinations are comparable across lakes, with means in the 50 to 70 mile range.

An important feature for some of these lakes is their varying water level within and between seasons. For jet skiers and motorized boaters water levels have important impacts on trip quality as some boat ramps become congested or unusable at low levels. For example, water levels at Boca and Stampede Reservoirs can change by 20-30 feet during a given summer. Furthermore, these changes can follow seasonal cycles (as has been typical for Boca), or inter-seasonal trends (as has been typical as for Stampede). Such pronounced changes in an important quality attribute for a given site make it naturally difficult for intercepted visitors to predict future trips. This supports the use of trip reports for past seasons to generate estimates for seasonal trip demand and economic benefits that apply to the latent population.

A more detailed picture of trip distributions for our sample is given in Table 2. The table depicts the mean over individuals of the number of trips taken to each of the five sites, distinguished by on-site versus off-site counts. For example, the first cell in the table indicates that the average number of trips to Boca Reservoir for those interviewed *at Boca* for the preceding season is 5.54. In contrast, respondents interviewed at other sites only took an average of 1.05 trips to Boca (second cell, first row). The "all" columns show the unweighted average of all trip counts for each site, regardless of on-or off-site status. The pronounced difference in trip averages between on- and off-site counts suggests that a correction for size-biased sampling may be indicated for past trip counts. Our estimation results confirm this postulation.

Estimation Results

We estimate two different models. Both are based on trip reports for the preceding season. Model 1 does not correct for on-site sampling, i.e. the avidity parameter α in (9) is constrained to zero in this model. Model 2 implements the avidity correction for past counts suggested in (9). It is important to emphasize that both models are based on the same underlying sample of visitors, i.e. the same data set except for the introduction of the transformation depending on α in model 2. A comparison of results generated by models 1 and 2 will illustrate the implications of ignoring avidity carryover.

Both models share the same basic set of explanatory variables, i.e. site and year-specific intercept terms to compactly capture site characteristics and potential inter-seasonal quality changes at each destination, a separate own-price term for each site[11], and the natural logarithm of income. The intercept terms correspond to the shifting vector $\mathbf{a_{ijt}}$ in (3), although we model these site indicators to be shared by all respondents for ease of estimation, i.e. $\mathbf{a_{ijt}} = \mathbf{a_{jt}}\ \forall i$. The models thus include all necessary components to estimate a Log II-type incomplete demand system (LaFrance and Hanemann, 1994).[12]

Estimation results are given in Table 3. The main result captured in the table is the location of the avidity parameter α near the center of its [0,1] support. In addition, this parameter is estimated with high precision as indicated by its comparatively small standard error. Invoking asymptotic normality, the 95% confidence interval for α is $\{0.57, 0.72\}$. At the same time, the value of the log-likelihood function at convergence is much improved by the introduction of the avidity parameter as shown in the last row of the table. A likelihood ratio test clearly rejects a null hypothesis of $\alpha = 0$.[13] As mentioned above, the likelihood function is not defined at $\alpha = 1$ for some individuals, which preempts the application of standard test procedures to verify this hypothesis. However, the tight confidence interval for α shown above suggests that it is highly unlikely that this parameter is located in the vicinity of one. We thus conclude that, at least for our application, (i) the distribution for past counts needs to be adjusted for on-site sampling (i.e. $\alpha \neq 0$), and (ii) that this adjustment is significantly different from the size-biased weight appropriate for current season counts (i.e. $\alpha \neq 1$).

The omission of this adjustment in model 1 translates into different estimates for slope coefficients as well as inflated variance terms compared to the corrected model. The difference in slope coefficients is especially pronounced for some of the price terms (Boca, Lahontan, Stampede). This, in turn, translates directly in substantial differences in trip and welfare predictions for these two models, as will be shown below. The price coefficient for Tahoe, which, counter-intuitively, emerges as positive in model 1 and insignificant in model 2, needs to be interpreted with caution. Given the large size of this lake and its multiple shoreline attractions, many intercepted jet skiers did not travel directly to the interview site, but launched their jet ski at a different location. This introduces measurement errors into the travel cost computations for such individuals.[14] We will thus exclude this site for trip predictions and welfare estimation.

Trip Predictions

Table 4 depicts model predictions for the number of seasonal trips to each site for the prototypical jet skier in the latent population. To compute these predictions we used the estimated parameters from our models in the expression for expected per-capita visits for the underlying population, given by δ_{ijt} in equation (7).[15] Specifically, for each model we take 10,000 draws from the empirical distribution of slope coefficients and variance terms. We then compute δ_{ijt} for each person and parameter draw and average over individuals. We then examine the properties of the resulting simulated distribution of mean trip predictions for each site. Table 4 reports the lower (LB) and upper bounds (UB) of the simulated 95% confidence interval for these trip means. In addition, we follow Moeltner (2003), Moeltner and Shonkwiler (2005), and Shonkwiler and Hanley (2003) by reporting a statistic that relates the point estimate of the mean to the spread of its confidence interval. This indicator of relative efficiency is denoted as "spread-over-mean" (s.o.m) in the table.

As can be seen from the table, trip predictions generated by model 1 are substantially larger than those produced by model 2. In fact, these predictions exceed even on-site sample counts for all sites, which casts serious doubt onto this model's ability to accurately recover latent population demand. Model 2, in contrast, produces latent trip predictions that lie below on-site sample counts for all sites, as one would expect. These predictions appear of plausible magnitude and are characterized by relatively tight confidence intervals and corresponding low s.o.m. values. It thus appears that past-season trip reports, if properly corrected for on-site sampling, provide a suitable basis to estimate latent population demand to the recreation system. On the other hand, the use of uncorrected past season trips leads to inflated estimates of population demand, as visitors with pronounced avidity for a given site are erroneously interpreted to behave like a typical user in the underlying population.

Welfare Estimates

For the same four sites, seasonal compensating variation (CV), which can be interpreted as "seasonal economic benefit" or "seasonal WTP to use the site", is captured in table 5. As shown in Shonkwiler (1999) for the Log II IDS seasonal CV for representative individual *i* and site *j* can be derived as

$$cv_{ij} = m_i - \left(m_i^{1-\beta_m} - \left(1-\beta_m\right) \cdot \left(\frac{1}{\beta_{p,j}} \cdot \exp\left(\alpha_j + \beta_{p,j} p_{ij} + 0.5 \cdot \sigma_{\varepsilon,j}^2 \right) \right) \right) \quad . \tag{11}$$

We simulate the distribution of the mean of (11) over individuals for each site in analogous fashion to the process described above for mean trip predictions.

From table 5 we observe a stark difference in seasonal benefit estimates between the two models. Specifically, welfare estimates generated by model 1 are two to six times higher than those produced by model 2. For example, model 1 predicts a welfare loss of \$336 for the typical jet skier if Lahontan Reservoir were to be closed or inaccessible for an entire season. The corresponding estimate for model 2 is substantially smaller at \$53 per person and season. Furthermore, confidence intervals are significantly tighter for the on-site corrected model, as indicated by lower s.o.m. statistics compared to model 1. As for trip predictions we infer that model 1 mis-interprets individual and site-specific avidity as a manifestation of “typical demand”, which naturally inflates resulting estimates of economic benefits.

IV) Conclusion

In this study we examine in more detail the statistical properties of on-site collected trip reports to a system of recreation sites for the preceding season. We find that for our sample of jet skiers visitation avidity for the site of intercept carries over across seasons. This requires a proper weighting of past season counts in the joint probability mass function of reported trips to avoid model mis-specification and biased estimation results. This weighting is different from the full size-biased weights appropriate for current-season counts.

There are several natural extensions to our model. First, an important question that remains to be examined is how models based on avidity-adjusted past trip counts compare to models using adjusted current season counts. If there are no pronounced changes in quality or accessibility across seasons one would expect both approaches to generate comparable predictions of trip demand and economic benefit. We abstain from such a comparison for our application, since much of our sampling took place in early- to mid season, and many respondents indicated substantial uncertainty when prompted to forecast trips for the remainder of the season[16]. An assessment of the magnitude and importance of forecasting errors would require a follow-up end-of-season polling of intercepted individuals, which was beyond the scope of our research project.

Second, an interesting future extension would be to examine at what rate avidity carryover erodes over time and if there exists individual heterogeneity in this respect. This would again require repeated observations for the same individuals over a longer time horizon, a notoriously difficult task in applications of recreation demand. Nonetheless, a better understanding of the dynamic nature of our avidity correction would be essential to generate more reliable forecasts for *future* trip demand based on on-site data.

Naturally, additional empirical research is needed to examine if and how avidity carryover changes over different recreational activities and destinations. However, based on the strong empirical evidence flowing from our application it appears to be prudent policy to ex ante allow for avidity corrections in any travel demand model based on past-season travel data collected on-site.

Notes:

[1] Naturally, asking visitors about past trips raises the issue of recall problems. We assume for this study that recall problems are expected to be considerably smaller than potential forecasting errors. The relative magnitude of these two types of measurement errors remains to be subjected to empirical examination.

[2] A richer inter-temporal model of consumer choice would link utilities across seasons, either through allowing for state dependence to directly enter seasonal utility (e.g. Adamowicz, 1994, Moeltner and Englin, 2004) or through inducing forward-looking rationality in a fully dynamic model (e.g. Provencher and Bishop, 1997). Implementation of the former approach requires substantially more choice occasions than were available for this analysis. The latter modeling strategy is computationally involved, especially given the econometric adjustments to site demands proposed in this study. In addition, as argued in Swait et al. (2004) consumers' recreation behavior is somewhat unlikely to flow from a fully dynamic optimization framework as mental processing costs would likely outweigh the gains in utility associated with (correctly) anticipating the effect of current decisions on future benefits.

[3] For simplicity, we assume travel costs and annual income to remain constant for the short time period (2 years) examined in our application.

[4] While these restrictions explicitly rule out cross-price effects in the uncompensated site-specific demand equations, they still allow for substitution between sites through compensated demands. Specifically, as shown in Englin et al. (1998) and Shonkwiler (1999) the Hicksian cross-price effects are non-zero as long as β_m is positive.

[5] The main intuition for allowing for site-specific variance in *heterogeneity effects* is that certain sites may have an unobserved quality attribute that is highly desirable to some and at the same

time perceived as a strong disamenity by others, while other sites may trigger less extreme responses in expected trip behavior. Similarly, the rationale for non-zero off-diagonal elements in $\mathbf{\Sigma}$ is that a given pair of sites may be similar or opposing with respect to an unobserved quality attribute. If *both* sites are relatively well endowed or relatively lacking in the attribute as perceived by the prototypical visitor, their covariance term is positive. If there is a distinct imbalance in the attribute across the two sites, their covariance is negative.

[6] We use 800 support points to simulate this sum. Estimation results stabilized at 600-800 points.

[7] The models are estimated using Matlab. The program code is available from the authors upon request.

[8] One of the destinations (Lake Topaz) was excluded from this study as very few visitors interviewed there took trips to any of the other sites in the system.

[9] The limitation of our sample to day trips is a utility-theoretic requirements, as longer trips constitute de facto a different commodity. Including trips of different length in the same sample would raise problems in interpreting aggregate welfare measures.

[10] Restrictions on survey length preempted collecting trip details for visits other than the one intercepted on-site. Our analysis thus rests on the implicit assumption that relevant trip details, such as length-of-stay, remain largely unchanged over all (past and future) trips for a given respondent.

[11] As in Moeltner and Shonkwiler (2005), we specify access price to incorporate a $0.3 per-mile driving cost for jet ski renters, and $0.4 for jet ski owners to allow for a "load penalty", as well as an opportunity cost of time-component computed as travel time in hours times 1/3 of the hourly wage rate.

[12] The survey also collected limited information on user characteristics, such as age, gender, and education level. However, none of these attributes emerged as significant in preliminary specifications.

[13] Since α is restricted under the alternative hypothesis, the LR statistic for this test follows a mixed χ^2 distribution, and standard LR test results may be biased towards not rejecting the null hypothesis (Chen and Cosslett, 1998). However, our LR-values are well above the upper bound for the critical χ^2 value for such a mixed distribution, and the adjustment procedure proposed by Chen and Cosslett (1998) would not affect our test results in this case.

[14] This problem did not become apparent until later in the survey period. As a result, the questionnaire did not capture the possibility of non-identical travel endpoints and interview sites.

[15] It should be noted that these population predictions cannot be directly compared to the sample statistics provided in table 2. Specifically, the "all" column in table 2 depicts a smeared average over all on- or off-site trips associated with a given lake. These sample averages are *not* indicative of latent user demand for the wider population.

[16] This is mainly due to the unpredictable water levels for some of the lakes, as noted earlier in the text.

References:

Adamowicz, W. L., Habit formation and variety seeking in a discrete choice model of recreation demand, Journal of Agricultural and Resource Economics 19 (1994) 19-31.

Aitchison, J., and C. H. Ho, The multivariate Poisson-lognormal distribution, Biometrika 76 (1989) 643-653.

Chen, Heng Z., and Stephen R. Cosslett, "Environmental quality preference and benefit estimation in Multinomial Probit models: A simulation approach," *American Journal of Agricultural Economics*, (1998), 512-520.

Egan, K., and J. Herriges, Multivariate count data regression models with individual panel data from an on-site sample, Journal of Environmental Economics and Management 52 (2006) 567-581.

Englin, J., P. Boxall, and D. Watson, Modeling recreation demand in a Poisson system of equations: An analysis of the impact of international exchange rates, American Journal of Agricultural Economics 80 (1998) 255-263.

Englin, J., D. Lambert, and W. D. Shaw, A structural equations approach to modeling consumptive recreation demand, Journal of Environmental Economics and Management 33 (1997) 33-43.

Englin, J., and J. S. Shonkwiler, Estimating social welfare using count data models: An application to the long-run recreation demand under conditions of endogenous stratification and truncation, The Review of Economics and Statistics 77 (1995) 104-12.

LaFrance, J. T., and W. M. Hanemann, On the integration of some common demand systems, Staff Papers in Economics 83-10, *Montana State University*, (1994).

LaFrance, J. T., and W. M. Hanemann, The dual structure of incomplete demand systems, American Journal of Agricultural Economics 71 (1989) 262-274.

Moeltner, K., Addressing aggregation bias in zonal recreation models, Journal of Environmental Economics and Management 45 (2003) 128-144.

Moeltner, K., and J. E. Englin, Choice behavior under time-variant quality: State dependence versus 'play-it-by-ear' in selecting ski resorts, Journal of Business & Economic Statistics 22 (2004) 214-224.

Moeltner, K., and J. S. Shonkwiler, Correcting for on-site sampling in random utility models, American Journal of Agricultural Economics 87 (2005) 273-535.

Patil, G. P., and C. R. Rao, Weighted distributions and size-biased sampling with applications to wildlife populations and human families, Biometrics 34 (1978) 179-189.

Provencher, B., and R. C. Bishop, An estimable dynamic model of recreation behavior with an application to Great Lakes angling, Journal of Environmental Economics and Management 33 (1997) 107-127.

Shaw, D., On-site samples' regression - problems of non-negative integers, truncation, and endogenous stratification, Journal of Econometrics 37 (1988) 211-223.

Shonkwiler, J. S., Systems of travel cost models of recreation demand, Paper presented at the Western regional research project (W-133) meetings in Pacific Grove, CA, 1995.

Shonkwiler, J. S., Recreation demand systems for multiple site count data travel cost models, in J. A. Herriges and C. L. Kling, (Eds.), Valuing recreation and the environment: Revealed preference methods in theory and practice, Edward Elgar, Cheltenham, U.K., and Northampton MA, 1999.

Shonkwiler, J. S., New estimators for the analysis of univariate and multivariate discrete data, Mimeo, Department of Resource Economics, University of Nevada, Reno (2001).

Shonkwiler, J.S., and N. Hanley, A new approach to random utility modeling using the Dirichlet Multinomial distribution, Environmental and Resource Economics 26 (2003) 401-416.

Swait, J., W. L. Adamowicz, and M. van Bueren, Choice and temporal welfare impacts: Incorporating history into discrete choice models, Journal of Environmental Economics and Management 47 (2004) 94-116.

Train, K., Discrete choice methods with simulation, Cambridge University Press, New York, 2003.

Train, K., Halton sequences for mixed logit, Working Paper, Department of Economics, University of California, Berkeley (1999).

Table 1: Sample Characteristics

Lake	Elevation (feet)	Surface area (square miles)	Shoreline (miles)	Distance (miles, one way) mean	min	max	On-site interviews	Past Season Trips
Boca	5700	1.5	15	49.1	7.2	221.9	31	308
Donner	5969	1.5	7.5	53.8	4.4	227.8	51	114
Lahontan	4150	23.2	69	74.4	1.1	205.6	33	346
Stampede	5949	5.4	25	56.8	8.7	228.0	26	105
Tahoe	6230	190.8	75	63.4	7.7	220.0	18	156
Total							159	1029

Table 2: Trip Statistics

Site	Past season trips (mean) on-site	off-site	all
Boca	5.58	1.05	1.94
Donner	1.37	0.41	0.72
Lahontan	8.21	0.60	2.18
Stampede	1.65	0.47	0.66
Tahoe	5.28	0.43	0.98
All sites	4.10	0.59	1.29

Table 3: Estimation Results

Variable	Model 1 Coeff.	s.e.		Model 2 Coeff.	s.e.	
Boca00	-5.861	(1.137)	***	-5.634	(0.977)	***
Donner00	-8.186	(1.177)	***	-8.028	(0.961)	***
Lahontan00	-2.992	(1.102)	***	-5.936	(0.732)	***
Stampede00	-6.702	(0.965)	***	-7.667	(1.234)	***
Tahoe00	-9.365	(1.101)	***	-8.375	(0.981)	***
Boca01	-6.503	(1.041)	***	-8.252	(1.041)	***
Donner01	-7.743	(1.131)	***	-8.363	(0.929)	***
Lahontan01	-3.569	(1.268)	***	-6.711	(0.729)	***
Stampede01	-8.026	(1.015)	***	-8.180	(1.086)	***
Tahoe01	-9.466	(1.032)	***	-8.599	(0.906)	***
price Boca	-0.060	(0.014)	***	-0.042	(0.007)	***
price Donner	-0.025	(0.007)	***	-0.021	(0.008)	**
price Lahontan	-0.091	(0.020)	***	-0.033	(0.002)	***
price Stampede	-0.042	(0.010)	***	-0.022	(0.007)	***
price Tahoe	0.004	(0.001)	***	0.002	(0.001)	
log(income)	0.580	(0.086)	***	0.494	(0.064)	***
log(1/α -1)	-	-	-	-0.578	(0.165)	***
α(avidity parameter)	-	-	-	0.641	(0.038)	***
Variances / Covariances						
Boca	7.862	(1.650)	***	7.664	(1.016)	***
Boca / Donner	3.247	(0.661)	***	2.606	(0.485)	***
Donner	5.574	(1.034)	***	4.709	(0.910)	***
Boca / Lahontan	4.056	(0.982)	***	3.125	(0.367)	***
Donner / Lahontan	-1.323	(0.610)	**	-0.170	(0.319)	
Lahontan	6.849	(2.251)	***	6.577	(0.417)	***
Boca / Stampede	4.617	(0.863)	***	3.314	(0.678)	***
Donner / Stampede	3.634	(0.688)	***	2.299	(0.439)	***
Lahnotan / Stampede	1.592	(1.114)		2.411	(0.673)	***
Stampede	5.969	(1.586)	***	4.873	(1.278)	***
Boca / Tahoe	3.305	(0.535)	***	3.586	(0.574)	***
Donner / Tahoe	3.266	(0.505)	***	2.292	(0.409)	***
Lahnotan / Tahoe	-0.396	(0.313)		-0.206	(0.333)	
Stampede / Tahoe	0.446	(0.315)		0.329	(0.421)	
Tahoe	5.807	(0.681)	***	3.276	(0.573)	***
Log-Lhf (abs. value)	829.865			797.485		

White-corrected standard error in parentheses

*=sign at 10%, ** = sign. at 5%, ***=sign. at 1%

Table 4: Seasonal Trip Predictions

Site		Model 1	Model 2
Boca	mean	14.91	5.17
	LB	5.66	3.30
	UB	36.17	8.29
	s.o.m.	2.05	0.96
Donner	mean	1.77	0.34
	LB	1.03	0.26
	UB	3.29	0.45
	s.o.m.	1.28	0.56
Lahontan	mean	28.61	1.74
	LB	4.61	1.34
	UB	123.47	2.28
	s.o.m.	4.15	0.54
Stampede	mean	2.25	0.46
	LB	1.07	0.30
	UB	5.21	0.76
	s.o.m.	1.83	0.99

Table 5: Seasonal Welfare Estimates

Site		Model 1	Model 2
Boca	mean	238.37	123.29
	LB	118.72	76.65
	UB	498.99	201.61
	s.o.m.	1.60	1.01
Donner	mean	76.79	17.85
	LB	37.42	9.94
	UB	162.27	34.85
	s.o.m.	1.63	1.40
Lahontan	mean	336.43	53.26
	LB	67.90	36.11
	UB	1335.77	81.10
	s.o.m.	3.77	0.84
Stampede	mean	52.79	29.15
	LB	32.45	12.36
	UB	96.36	78.95
	s.o.m.	1.21	2.28